R E P O R T R E S U M E S

ED 016 653 TE 000 189

CLASS SIZE AND ACHIEVEMENT GAINS IN SEVENTH- AND EIGHTH-GRADE ENGLISH AND MATHEMATICS.

BY- JOHNSON, MAURITZ SCRIVEN, ELDON

PUB DATE 67

EDRS PRICE MF-$0.25 HC-$0.56 12P.

DESCRIPTORS- *ACHIEVEMENT GAINS, *CLASS SIZE, *ENGLISH INSTRUCTION, *HOMOGENEOUS GROUPING, *MATHEMATICS INSTRUCTION, GRADE 7, GRADE 8, READING COMPREHENSION, TEACHING LOAD, ARITHMETIC, NEW YORK STATE

USING DATA OBTAINED BY THE QUALITY MEASUREMENT PROJECT ON SOME 70,000 PUPILS, THIS STUDY ATTEMPTED TO PRODUCE EVIDENCE REGARDING THE INFLUENCE OF CLASS SIZE AND CLASS HOMOGENEITY ON ACHIEVEMENT GAINS IN GRADES 7 AND 8. A TOTAL OF 130 ENGLISH AND 135 MATHEMATICS CLASSES CLASSIFIED ACCORDING TO SIZE AND HOMOGENEITY WERE EXAMINED. THE READING COMPREHENSION AND ARITHMETIC TEST SCORES ON THE IOWA TEST OF BASIC SKILLS WERE USED AS THE MEASURES OF ACHIEVEMENT. RESULTS INDICATED THAT GAIN DIFFERENCES IN RESPECT TO CLASS SIZE AND CLASS VARIABILITY WERE GENERALLY VERY SMALL AND INCONSISTENT. BECAUSE TWO-THIRDS OF THE CLASSES STUDIED CONSISTED OF FROM 23 TO 32 PUPILS, THE LARGEST AND SMALLEST CLASSES (LARGER THAN 34 AND SMALLER THAN 24 STUDENTS) WERE ISOLATED FOR SEPARATE COMPARISON. RESULTS CONFIRMED THAT THERE WAS NO INSTANCE OF A SIGNIFICANT DIFFERENCE IN ACHIEVEMENT GAIN EVEN BETWEEN THESE EXTREME GROUPS. ALTHOUGH THESE TESTS DO NOT MEASURE ALL TYPES OF ACHIEVEMENT, THEY DO SUGGEST THAT ATTENTION MIGHT MORE PROFITABLY BE DIRECTED TOWARD REDUCING THE NUMBER OF CLASSES ASSIGNED TO ONE TEACHER THAN TOWARD REDUCING THE SIZE OF THE CLASSES THEMSELVES. (THIS ARTICLE APPEARED IN "THE SCHOOL REVIEW," VOL. 75, NO. 3, AUTUMN 1967.) (DL)

UNIVERSITY OF CHICAGO

The School Review

A QUARTERLY

Published by THE UNIVERSITY OF CHICAGO PRESS for the
DEPARTMENT OF EDUCATION OF THE UNIVERSITY OF CHICAGO

AUTUMN 1967

MAURITZ JOHNSON
Cornell University
ELDON SCRIVEN
Northern Illinois University

Class Size and Achievement Gains in Seventh- and Eighth-Grade English and Mathematics

"For virtually the entire history of modern secondary education in this Nation . . . ," a report of the U.S. Office of Education notes, "there has been endless discussion over the proper number of pupils to be assigned a given secondary school class."[1] The common assumption has been that students achieve better in small classes. This study examines some evidence bearing on that assumption.

Current proposals by Trump and Baynham[2] and others tend to change the nature of the problem. Schools are urged to make flexible use of classes of various sizes for different kinds of learning activity. The effects of such variation on achievement will merit study, but the data presented here do not concern this practice. They deal with the conventional, and still most prevalent, arrangement whereby all classes are somewhat similar in size, are taught by one teacher, and do not vary in composition from day to day.

Obviously, many factors determine the sizes of instructional groups in schools. The nature of the subject, the ability levels of groups, the capacities of available classrooms, and relative costs must all be taken into consideration. In terms of the additional teachers and classrooms that would be needed, a reduction of average class size by even one pupil throughout the nation would be enormously expensive. There may be many values of smaller classes which would amply justify the increased costs involved. But to the

extent that better academic achievement is one of the expected benefits, it is important to know what relation, if any, there is between class size and achievement.

Average class size in junior high schools has been reported[3] to be about 31. An oft-recommended goal, which accords with teachers' preferences,[4] is a class size of 25. Thus, it is within this narrow range of 25–30 that the controversy centers. The average size of classes considered in the study reported here was between 28 and 29.

In 1964, Millman and Johnson[5] examined the achievement gains of a large number of instructional groups in terms of their homogeneity with respect to initial achievement level. Further analysis of the same data now furnishes some evidence regarding the influence of class size on achievement gains.

PROCEDURE

Data obtained by the Quality Measurement Project and made available by the New York State Education Department include the scores of large numbers of seventh- and eighth-grade students tested successively in 1957–58 and 1958–59 with the Iowa Basic Skills battery. Some 70,000 pupils in approximately a hundred representative school systems were tested. Class lists were obtained for thirty of these systems. Out of these were selected only those classes which were tested approximately a year apart and sufficiently early or late in the year so that the gain could appropriately be identified with either the seventh or eighth grade; only those classes in which the pupils still available for the second testing had mean initial scores within three-tenths of a grade of the mean initial score of the total class; and only those classes the true size of which was definitely known. A total of 130 English classes and 135 mathematics classes were available for study, 103 in grade 7 and 162 in grade 8. In many instances, as a result of conventional scheduling practices, pupils comprising an English class were the same ones who made up a mathematics class. Because this was not always the case, however, and

because certain sections did not meet the stated criteria, the two subjects, as well as the two grades, were treated separately.

The ITBS reading comprehension test scores were used as the measure of the dependent variable for the English classes, and for the mathematics classes, the scores on the arithmetic test. Gain scores represent the difference between the grade equivalent achieved in the seventh and the eighth grades. Classes were classified as homogeneous or heterogeneous on the basis of the standard

TABLE 1

CUTTING POINTS USED TO CLASSIFY CLASS SECTIONS ON INITIAL ACHIEVEMENT LEVEL, VARIABILITY, AND SIZE (TEST DATA IN GRADE EQUIVALENTS ON IOWA TESTS OF BASIC SKILLS)

Classification	English Classes		Mathematics Classes	
	Grade 7 ($N=48$)	Grade 8 ($N=82$)	Grade 7 ($N=55$)	Grade 8 ($N=80$)
Initial achievement level (section mean):				
Above average	7.851 and over	8.900 and over	7.551 and over	8.300 and over
Below average	Below 7.851	Below 8.900	Below 7.551	Below 8.300
Class variability (standard deviation):				
Heterogeneous	1.045 and over	1.231 and over	0.651 and over	0.771 and over
Homogeneous	Below 1.045	Below 1.231	Below 0.651	Below 0.771
Class size (number of pupils):				
Large	29 and over	28 and over	29 and over	28 and over
Small	Below 29	Below 28	Below 29	Below 28

deviation of students' scores in the initial (seventh-grade) testing; they were classified by level as high or low on the basis of the mean initial scores. The cutting points in the triple dichotomization by size, variability, and level are given in Table 1.

RESULTS

The partitioning created eight categories of classes for each subject in each grade, a total of thirty-two. The mean gain of each class was computed. The means of the mean gains for all classes in each of the categories are given in Table 2. The number of classes in each category is shown.

TABLE 2

MEAN GAINS ON IOWA TESTS OF BASIC SKILLS ACHIEVED BY CLASSES CATEGORIZED BY INITIAL ACHIEVEMENT LEVEL, VARIABILITY, AND SIZE (DATA IN GRADE EQUIVALENTS)

Initial Achievement Level	Small			Large			Total		
	Homogeneous	Heterogeneous	Total Small	Homogeneous	Heterogeneous	Total Large	Homogeneous	Heterogeneous	Total Classes
Seventh-Grade English Classes (Reading Scores)									
Above average:									
N	7	5	12	6	6	12	13	11	24
Mean	0.82	0.89	0.85	0.66	0.47	0.57	0.75	0.66	0.71
Below average:									
N	6	5	11	5	8	13	11	13	24
Mean	0.71	0.66	0.69	0.67	0.89	0.80	0.70	0.80	0.75
Total:									
N	13	10	23	11	14	25	24	24	48
Mean	0.77	0.77	0.77	0.67	0.71	0.69	0.72	0.74	0.73
Eighth-Grade English Classes (Reading Scores)									
Above average:									
N	11	12	23	9	9	18	20	21	41
Mean	0.82	0.76	0.79	0.71	0.80	0.75	0.77	0.78	0.78
Below average:									
N	12	6	18	9	14	23	21	20	41
Mean	0.80	0.69	0.77	0.86	0.87	0.87	0.83	0.82	0.82
Total:									
N	23	18	41	18	23	41	41	41	82
Mean	0.81	0.74	0.78	0.78	0.84	0.82	0.80	0.80	0.80
Seventh-Grade Mathematics Classes (Arithmetic Scores)									
Above average:									
N	8	5	13	5	10	15	13	15	28
Mean	1.35	1.43	1.38	1.01	0.94	0.96	1.22	1.10	1.16
Below average:									
N	9	6	15	6	6	12	15	12	27
Mean	0.77	1.03	0.87	0.96	1.13	1.05	0.85	1.08	0.95
Total:									
N	17	11	28	11	16	27	28	27	55
Mean	1.04	1.21	1.11	0.99	1.01	1.00	1.02	1.09	1.06
Eighth-Grade Mathematics Classes (Arithmetic Scores)									
Above average:									
N	10	13	23	12	6	18	22	19	41
Mean	1.53	1.52	1.52	1.76	1.71	1.75	1.66	1.58	1.62
Below average:									
N	11	5	16	7	16	23	18	21	39
Mean	0.98	1.04	1.00	1.19	1.16	1.17	1.06	1.13	1.10
Total:									
N	21	18	39	19	22	41	40	40	80
Mean	1.24	1.39	1.31	1.55	1.31	1.42	1.39	1.34	1.36

It is evident from Table 3 that the gain differences between large and small classes are generally very small and inconsistent. There are grade and subject differences as well as those among levels and variability. For both English and mathematics, however, seventh-grade gains were greatest for small heterogeneous classes, while in

TABLE 3

GAIN SCORE DIFFERENCES FOR MAIN EFFECTS

Subject and Grade	Class Size		
	Small	Large	Difference
English (7)	0.77	0.69	+0.08
English (8)	0.78	0.82	−0.04
Math (7)	1.11	1.00	+0.11
Math (8)	1.31	1.41	−0.10
	Class Variability		
	Homogeneous	Heterogeneous	Difference
English (7)	0.72	0.74	−0.02
English (8)	0.80	0.80	0.00
Math (7)	1.02	1.09	−0.07
Math (8)	1.39	1.34	+0.05
	Initial Achievement Level		
	Above Average	Below Average	Difference
English (7)	0.71	0.75	−0.04
English (8)	0.78	0.82	−0.04
Math (7)	1.16	0.95	+0.21
Math (8)	1.62	1.10	+0.52

the eighth grade they were greatest for large homogeneous classes. In both grades, below-average classes showed greater gains than above-average classes on the reading test, while on the arithmetic test the gains were more in line with expectations. It is possible that there was a "topping" effect on the reading test or that schools tended to try more deliberately to improve pupils' reading abilities in lower-ability classes.

The first-order interaction effects involving class size are indicated in Table 4. In ten of the sixteen situations, gains were greatest for the larger groups. The summary in Table 5 shows that large groups had the advantage seven out of eight times for all eighth-grade classes, seven out of eight times for below-average classes in both grades, six out of eight times for all mathematics classes, and six out of eight times for heterogeneous classes in both subjects.

TABLE 4

GAIN SCORE DIFFERENCES FOR FIRST-ORDER INTERACTION EFFECTS INVOLVING CLASS SIZE

Category	Gains		
	Small	Large	Difference
Above average homogeneous:			
English (7)	0.82	0.66	+0.16
English (8)	0.82	0.71	+0.11
Math (7)	1.35	1.01	+0.34
Math (8)	1.53	1.76	−0.23
Above average heterogeneous:			
English (7)	0.89	0.47	+0.42
English (8)	0.76	0.80	−0.04
Math (7)	1.43	0.94	+0.49
Math (8)	1.52	1.71	−0.19
Below average homogeneous:			
English (7)	0.71	0.67	+0.04
English (8)	0.80	0.86	−0.06
Math (7)	0.77	0.96	−0.19
Math (8)	0.98	1.19	−0.21
Below average heterogeneous:			
English (7)	0.66	0.67	−0.01
English (8)	0.69	0.87	−0.18
Math (7)	1.03	1.13	−0.10
Math (8)	1.04	1.16	−0.12

Because so many interaction effects seemed to be present, a triple classification analysis of variance was carried out for both subjects in both grades. Since cell sizes were not equal, a table of random numbers was used to eliminate class sections whenever necessary to bring the number in each cell to five. A summary of the analysis is given in Table 6.

It will be noted that only five effects proved to be statistically significant. None of the others was significant at even the .05 level.

TABLE 5

SUMMARY OF RELATION OF CLASS SIZE ON ACHIEVEMENT GAINS FOR VARIOUS CATEGORIES

Categories	Situations Favoring	
	Small	Large
Subject:		
English	4	4
Mathematics	2	6
Grade:		
7	5	3
8	1	7
Level:		
Above average	5	3
Below average	1	7
Variability:		
Homogeneous	4	4
Heterogeneous	2	6

TABLE 6

SUMMARY OF TRIPLE CLASSIFICATION ANALYSIS OF VARIANCE

Source	English						
	Grade 7				Grade 8		
	df	Sum of Squares	Mean Square	*F*	Sum of Squares	Mean Square	*F*
Class size	1	8.99	8.99	1.54	8.58	8.58	1.29
Variability	1	0.63	0.63	0.11	0.94	0.94	0.14
Achievement level	1	1.35	1.35	0.23	4.92	4.92	0.74
Size×variability	1	0.24	0.24	0.04	6.23	6.23	0.94
Size×level	1	66.72	66.72	11.46*	3.13	3.13	0.47
Variability×level	1	14.74	14.74	2.53	4.46	4.46	0.67
Size×variability×level	1	14.77	14.77	2.54	0.43	0.43	0.06
Within	32	186.17	5.82	.....	265.85	6.65	.....
Total[a]	39	293.61	.....	.....	294.54	.....	.....
	Mathematics						
Class size	1	15.90	15.90	2.99	19.83	19.83	1.99
Variability	1	15.28	15.28	2.87	0.56	0.56	0.06
Achievement level	1	88.21	88.21	16.58**	245.97	245.97	24.67**
Size×variability	1	4.08	4.08	0.77	8.35	8.35	0.84
Size×level	1	138.01	138.01	25.94	7.32	7.32	0.73
Variability×level	1	42.60	42.60	8.01*	2.64	2.64	0.26
Size×variability×level	1	2.07	2.07	0.39	4.01	4.01	0.40
Within	32	170.24	5.32	.....	310.59	9.97	.....
Total	39	476.39	.....	.....	599.27	.....	.....

[a] Analysis for English, Grade 8, based on 48 classes; *df.* within = 40.

* Significant at .01 level.

** Significant at .001 level.

Achievement level is a significant determiner of gains in mathematics in both grades. Class size interacts with achievement level in both subjects, but only in the seventh grade.

Although the classes studied varied in size from 10 to 36, two-thirds of them consisted of from 23 to 32 pupils. It may be argued

TABLE 7

MEDIAN ACHIEVEMENT GAINS OF THE LARGEST AND SMALLEST CLASSES (GAINS IN GRADE EQUIVALENTS)

Subject, Grade, Type	Small			Large			Gain Diff.	Mann-Whitney *U*
	N	Med. Size	Med. Gain	*N*	Med. Size	Med. Gain		
English (7):								
Total	6	18.5	0.70	11	35.0	0.53	+0.17	25 (13)*
Above average	1	24.0	0.58	6	35.0	0.49	+0.09	1 (0)
Below average	5	18.0	0.76	5	35.0	0.80	−0.04	15 (3)
Homogeneous	3	19.0	0.76	4	34.5	0.62	+0.14	5 (0)
Heterogeneous	3	18.0	0.64	7	35.0	0.45	+0.19	7 (1)
English (8):								
Total	16	22.0	0.77	8	35.0	0.85	−0.08	62 (31)
Above average	5	22.0	0.95	0	.....	.....	.....	.....
Below average	11	22.0	0.70	8	35.0	0.85	−0.15	35 (19)
Homogeneous	12	22.0	0.83	1	35.0	0.85	−0.02	6 (0)
Heterogeneous	4	22.0	0.70	7	35.0	0.85	−0.15	16 (3)
Math (7):								
Total	8	19.5	0.77	13	35.0	0.94	−0.17	75 (24)
Above average	2	24.0	1.05	7	35.0	0.85	+0.20	5 (0)
Below average	6	18.5	0.68	6	35.0	1.02	−0.34	31 (5)
Homogeneous	5	18.0	0.71	4	34.5	0.97	−0.26	5 (1)
Heterogeneous	3	24.0	1.14	9	35.0	0.94	+0.20	12 (2)
Math (8):								
Total	11	21.0	1.04	9	35.0	1.11	−0.07	45 (23)
Above average	3	22.0	1.54	0	.....	.....	.....	.....
Below average	8	21.0	0.96	9	35.0	1.11	−0.15	53 (15)
Homogeneous	6	21.5	1.04	1	34.0	1.39	−0.35	1 (0)
Heterogeneous	5	21.0	1.04	8	35.0	1.08	−0.04	22 (6)

* Figure in parentheses is critical value of *U* at .05 level for two-tailed test. Probability of any larger *U* exceeds .05; hence none of differences shown is significant at that level.

that a comparison of more extreme sizes would reveal a relationship that is obscured in the above analysis. To test this possibility, the largest and smallest classes were isolated for separate comparison. None of the largest classes contained fewer than 34 students, and none of the smallest contained more than 24. Their median sizes and median achievement gains are given in Table 7.

A Mann-Whitney *U*-test[6] was performed on the distribution for each subject in both grades. The results confirmed the general finding of the study: There was no instance of a significant difference in achievement gain even between these extreme groups. The medians for subgroupings classified by initial achievement level and variability are reported merely to indicate how negligible the differences were. In most instances, however, the number of cases in these categories is so small that it would be impossible for a significant difference to manifest itself.

CONCLUSIONS AND DISCUSSION

An examination of the achievement gains made by some 7,500 seventh- and eighth-grade pupils in 265 English and mathematics classes reveals no consistent effect of class size on the gains. Of course these tests do not measure all of the types of achievement toward which instruction was directed, nor do they reveal other benefits which might accrue to students from being in small classes. But insofar as improvement on the Iowa Test of Basic Skills is concerned, it does not seem to make any difference how large the classes are, *within the limits included in this sample.* It should not be concluded, of course, that the same results would be obtained if class size were allowed to go as high as one hundred or two hundred. Neither do the findings rule out the possibility that higher achievement gains might, with different procedures, be attained by smaller classes. All that the present data indicate is that, among the classes studied, such superior attainment was not in evidence.

Inasmuch as essentially the same data were used in this study as in Millman and Johnson's,[7] again no consistent relation is revealed between class variability and achievement gains. Oddly, it is only for mathematics that even initial achievement level is consistently related to gains.

Considering the small differences entailed and the interaction of numerous factors, it does not mean much to conclude from this study

that large heterogeneous mathematics sections should be advocated for below-average eighth-grade students. Nevertheless, the results suggest that uncritical worship of small classes for all subjects, grades, and ability levels is unjustified.

It is important to distinguish between class size and teachers' loads. The total number of contacts may be more significant than the number at any one time. Four classes of thirty may be preferable to five of twenty-four. The purpose of reducing class size must also be considered. Just because reading scores are not improved by reducing the sizes of English classes does not mean that writing ability will not thereby be improved.

Secondary school teachers need more time for scholarly activities. If their schedules called for fewer but larger classes, some time might be gained for this purpose. It is also conceivable that in some instances generally larger classes and total loads, with appropriate additional remuneration, might both be economically advantageous and relieve teacher shortages in critical fields. Experimentation conducted free of the prevailing bias against large classes is badly needed. The results of the present study tend to undermine one of the bases of the bias.

It may be noted also that the results apparently accord with recent findings in Sweden. Torsten Husén[8] has reported a study by Sixten Marklund[9] that bore upon a proposal in the Parliament to "increase the educational productivity" of the country by "a successive reduction in class size." Husén states that "Marklund was able to make it quite clear that within the range under consideration there was no evidence to support such a prediction."

NOTES

1. Edmund A. Ford and Virgil R. Walker, *Public Secondary Schools* (Statistics of Education in the United States, No. 1 [Washington: U.S. Office of Education, 1961]), p. 15.

2. J. Lloyd Trump and Dorsey Baynham, *Guide to Better Schools: Focus on Change* (Chicago: Rand McNally & Co., 1961).

3. "Class Size in Secondary Schools," *NEA Research Bulletin*, XLIII (February, 1965), 21.

4. Ellsworth Tompkins, *What Teachers Say about Class Size* (Washington: U.S. Office of Education, Circular 311, 1949), p. 4.

5. Jason Millman and Mauritz Johnson, "Relation of Section Variance to Achievement Gains in English and Mathematics in Grades 7 and 8," *American Educational Reseach Journal*, I (February, 1964), 47–51.

6. Sidney Siegel, *Non-parametric Statistics* (New York: McGraw-Hill Book Co., 1956), p. 116.

7. Millman and Johnson, *op. cit.*

8. Torsten Husén, "A Case Study in Policy-oriented Research: The Swedish School Reforms," *School Review*, LXXIII (Autumn, 1965), 206–25.

9. Sixten Marklund, *Skolklassens Storlek och Structur* (Stockholm: Almqvist & Wiksell, 1962).